

Supplemental Material

Variability in the omics

Sources of Technical Variability

Transcriptomics: The complex nature of microarray experiments, with many individual manipulations between biological sampling and data interpretation, make technical variability a serious concern. The process of interpretation is especially vulnerable to technical variability because of the simultaneous measurement of hundreds or thousands of endpoints and these are compounded by the high cost of the technique which has often led to the use of suboptimal sample sizes.

Technical variation in transcriptomics measurements can arise from many sources, including sampling processes, non-ideal RNA isolation and storage, variations in RNA labeling (including lability of probes), imperfect hybridization and subsequent data analysis methodologies. Transcriptomics has been the subject of quite extensive evaluations of technical variation (reviewed in detail in (de Koning et al. 2007; Fuscoe, et al. 2007; Mattes 2008; Thompson and Hackett 2008; Walker and Hughes 2008). Several consortia have recently examined technical variability within laboratories, between laboratories and between different expression monitoring platforms (Bammler et al. 2005; Kuo et al. 2006; Shi et al. 2006; Beyer et al. 2007; Chen et al. 2007; Arikawa et al.

2008; Fielden et al. 2008; Kohlmann et al. 2008). These large scale studies have demonstrated that microarray analysis can be performed with good intra- and inter-laboratory reproducibility for genes that have copy numbers of at least 5 to 10-fold above their detection limits and this has been confirmed by alternative methods such as quantitative real time PCR (Canales et al. 2006).

Several key areas have been identified that can cause technical artefacts. These include study design, inadequate numbers of samples, and methods for sample acquisition, preparation, storage, processing and analysis. Once capture of the biological target has been achieved it is important that equivalent portions of tissue are dissected because of the likelihood of cellular and functional heterogeneity of the tissue and zonation of gene expression (Gebhardt 1992; Oinonen and Lindros 1998; Boedigheimer et al. 2008). It is important to recognize that batch effect as a source of error is likely and that biological samples are processed at random in each batch rather than each batch analyzing biological material from a single sampling event. Sample integrity is critical to successful analysis. Tissue thawing can lead to a rapid loss of RNA integrity thereby dramatically affecting array quality especially if probes are located more than 1000 bp from the transcriptional end of the mRNA (Thompson et al. 2007). Standardized protocols and external controls for quality control must be established to assure reliability. One aspect that can contribute to inter-laboratory studies is the use of external RNA controls (Baker et al, 2005). The quality of microarray analysis can be greatly improved by inclusion of adequate documentation that capture important aspects of experimental variables such as

treatments, sample quality and methods, therefore helping to assess microarray quality issues due to protocol variation.

Design of microarrays can introduce technical variation especially if multiple platforms are used for analysis. Disagreement in results from different platforms can result from probes hybridizing with different efficiencies to non-overlapping target sequences and in the case of cDNA probes, this may be compounded by cross hybridization with multigene family members. The format of the microarrays can affect variability with commercial products having better quality control and reproducibility with different sample labeling protocols contributing negligible variations in results (Kuo et al., 2006, Patterson et al., 2006). While different formats give similar results in terms of expression patterns, variation in signal intensity and resulting expression values can be obtained between different laboratories and microarray formats (Chen et al. 2007).

Perhaps the most important technical variable is in the normalization and statistical interpretation of array data. Preliminary steps should be employed to eliminate outliers (excessive chip-to-chip variation) that may be as high as 5-10% (Boedigheimer et al. 2008; Hershey et al. 2008). Normalization is essential to eliminate the effects of variable cDNA labeling but different normalization protocols and statistical analysis algorithms generate different lists of differentially expressed genes. For chip-based arrays, normalization by robust multichip average (RMA) and gene-chip RMA, are common approaches and may have advantages over traditional normalization with housekeeping genes (Irizarry et al. 2003; Hershey et al. 2008), especially since in the present context

toxicants are being studied and by definition these will adversely affect cellular architecture and basal metabolic processes. While some groups advocate simple cut off rules based on fold-change versus controls and significance tests (Shi et al. 2006), this can lead to high false positive errors if a no-treatment effect is not considered (Chen et al. 2007). Reproducibility can be maximized if analyses are based on biological functions such as defined by gene ontology terms and toxicological function (Bammmler et al. 2005; Fielden et al. 2008). Analysis of species of ecological interest is complicated by lack of genome sequence and poor annotation.

Proteomics: Global analyses of the proteome are now becoming possible with new technological innovations, however, sample preparation and manipulations can be exceedingly complex and thus variability is of considerable concern. Recently there have been a few attempts to ascertain intra-individual variability but these have not, so far, been extended to inter-laboratory comparisons, possibly due to the lack of established standards for representing proteomic data.

Inter-experimental reproducibility is not good using 2-dimensional polyacrylamide gel electrophoresis approaches, however, an excellent differential display technique (DIGE) allows multiple samples to be compared on the same gel and is a powerful tool in biomarker discovery for laboratory exposure studies which is quantifiable. Potentially useful biomarker candidates must be carefully selected from proteins that are not subject to sexual, nutritional and naturally high variations in expression, moreover, they must be sufficiently abundant and resolvable from other proteins for reliable quantitation. Recent studies indicate that a sample size of some 7 replicates creates reliable data. There is

obviously a pressing need for establishment of standardized selection procedures. Whilst proteins of genomically characterized animals can be identified relatively easily by mass spectrometry of excised proteins using MALDI-TOF instruments, protein identification in non-characterized organisms requires *de novo* sequencing methods, precluding routine use until sufficiently comprehensive databases have been assembled.

Increasingly, proteomic techniques use liquid-chromatography (LC) separations coupled with electrospray ionization (ESI) MS and tandem mass spectrometry (MS/MS) for the characterization of the separated peptides or proteins. Sample preparation procedures are potentially much less variable and theoretically they should be able to analyze a larger proportion of the proteome, however, extensive sequence databases of the study organism are required and methods for accurate inter-individual quantitation are lacking and still under development.

Metabolomics. The ultimate expression of an organisms' phenotype is the profile of metabolites in its cells and bodily fluids. Often their existence is transient as many are rapidly biotransformed therefore sampling, sample preservation and extraction techniques are critical for high reproducibility. Standardisation of methods is therefore essential and progress is being made on this front.

Keun et al. (2002) assessed the analytical reproducibility of an NMR metabolomics experiment by analysing two identical sets of rat urine samples from an acute toxicity study. The analyses were performed at two sites and principal components analyses

(PCA) revealed extremely similar descriptions of the metabolic responses to hydrazine. In one study (Bertram et al. 2007), identical sets of human urine before and after dietary intervention were measured using 250, 400, 500 and 800 MHz NMR spectrometers. When analysed by partial least squares discriminant analysis (PLS-DA), the loadings were found to comprise of the same spectral regions implying that the same metabolites were discriminating pre- and post-dietary intervention, independent of magnetic field strength. Most recently, an intercomparison exercise involving seven laboratories evaluated the accuracy, precision and efficacy of ^1H NMR metabolomics for environmental research (Viant et al. 2009).

The study comprised the analysis of both synthetic metabolite mixtures as well as European flounder (*Platichthys flesus*) liver extracts from clean and contaminated sites. For both sample types, PCA revealed highly similar scores plots across all laboratories. Furthermore, the same metabolic biomarkers that discriminated fish from clean and contaminated sites were discovered by all the laboratories. Taken together, these studies clearly demonstrate that NMR-based metabolomics can generate data that are sufficiently reproducible between laboratories to support its use in regulatory studies. No such intercomparison exercises have yet to be reported for mass spectrometry based metabolomics studies.

Sources of biological variability

Biological variation can be conveniently sub-categorised into variation within the control group or population (i.e. intra-class variation that is often unrelated to the toxic stressor being studied) and variation between the control and exposed groups (i.e. inter-class variation). In general, experiments should be designed to minimise both technical and intra-class variation, thereby maximising inter-class differences that can be explored using data mining techniques. Meaningful results depend on technical variability being less than biological variability.

As a prelude to this discussion, two distinct scenarios must be considered. In chemical testing (e.g. OECD) a limited number of model organisms are used worldwide for controlled laboratory studies. In environmental monitoring, locally relevant sentinel species, with little supporting genomics information, are typically used and sampled directly from the environment. These two scenarios will be associated with significantly different degrees of biological variation.

Variability between individuals within a given population is an essential component of population health and sustainability as it encodes for phenotypic flexibility and ability to acclimate to changing conditions and is the vehicle of evolution. Environmental stressors can impact this and affect the phenotypic variation between individuals. Reduced variability is potentially adverse to the sustainability of the population. Little is understood about the stressor-induced changes on these parameters in an environmentally relevant context. Some examples in the literature indicate that individuals within a population may have different degrees of susceptibility to estrogenic exposure, resulting

in a large spread in the degree of response at low concentrations of estrogens and a more consistent response between individuals at high concentrations (Thorpe et al. 2001; Thorpe et al. 2003). The implications of the decreased variability caused by chemical exposure for the sustainability of wild fish populations are unknown. Potentially, if such an effect was also occurring in the wild, it could impact on the ability of the population to cope with changing environmental conditions and compromise its sustainability. It is clear that individual variation might be considered as an endpoint in toxicological studies when population level effects are being addressed as chemicals may impact on this parameter and compromise the ability of populations to survive in environments constantly under pressure by changes in natural factors and anthropogenic stressors.

Inter-individual biological variation can be associated with two levels of organisation, generally defined as genotypic and phenotypic variation. Genetic variation is inherent in individuals within populations and accounts for the phenotypic plasticity allowing individuals to acclimate and populations to adapt to changing conditions. Genetic variation within a group of organisms is apparent at several levels, including polymorphisms, copy-number variants, alternative splicing, post-transcriptional and post-translational regulation and epigenetic modifications. Furthermore, organisms may vary in strain or clonal line. In general, genetic variability will be considerably higher (and less well characterised) for environmentally sampled organisms compared to model organisms unless out-bred colonies have been maintained. Furthermore the identification of closely related species (eg *Mytilus spp.*) is sometimes difficult for environmentally sampled organisms, which can be a major source of error and a confounding issue in the

interpretation of ecotoxicogenomics data. Phenotypic variation is strongly influenced by the environment and its interaction with the unique physiological conditions associated with individual organisms. This form of variation can arise from factors such as age, stage of life cycle and reproductive cycle, sex, nutritional status and general health. Phenotypic variation is also likely to be lower in studies of model organisms under laboratory conditions where it is practical to use standardized experimental conditions. Many of these factors cannot be controlled for environmentally-sampled animals, but recommendations applied to current biological-effects monitoring regarding sampling and documentation (by EPA, ICES etc) should be adhered to. The NERC Environmental Bioinformatics Centre (NEBC) has recommended those parameters that should be recorded and reported for transcriptomic submissions (envgen.nox.ac.uk/posters/MIAME_Env.ppt) as has the Environmental Context working subgroup of the Metabolomics Standards Initiative for a metabolomics study (Morrison et al., 2007). While it might be anticipated that model organisms raised in a controlled laboratory exhibit less environment- and capture-induced variation than similar animals living in the wild, the situation is far less clear for wild animals housed within the laboratory for short periods. Hines et al. (2007) showed that direct sampling of *Mytilus galloprovincialis*, a marine mussel, from the environment resulted in less metabolic variation than for animals from the same cohort that were returned to the laboratory in an attempt to allow their metabolome to equilibrate in controlled conditions.

Mammalian toxicogenomics studies have provided valuable information on baseline fluctuations in gene expression due to study condition and/or endogenous factors. A

consortial effort was recently undertaken by the HESI Genomics Technical Committee to examine microarray data from control animals from toxicogenomics studies of rat liver and kidney (Boedigheimer et al. 2008). Gender, organ section, fasting state, and strain emerged as study factors that contributed highly to variability in gene expression, whereas other factors, such as age, vehicle administration route, sacrifice method, and dose frequency were not major contributors to baseline variance. Genes with high variability, identified in the study, include many of interest in toxicology, such as those involved in xenobiotic metabolism, androgen and estrogen metabolism, steroid biosynthesis, and antigen processing and presentation. Low variance genes were also identified, and included those involved in protein metabolism and immune response. Such low variance genes may prove valuable as study controls. This collaborative effort determined the impacts of key study factors on measured gene expression in a toxicogenomics study and illustrated the importance of defining the baseline gene expression against which stressor-induced gene expression changes are to be evaluated.

Others have considered expression level variation as a genetic Quantitative Trait that can be used to identify loci that regulate gene expression and thus explain the mechanisms behind the variation (Williams et al. 2007). While it is possible to establish the importance of genetic factors in dictating mRNA expression level, variation between independent studies frequently generates conflicting data which have sometimes been attributed to a technical failure of array technology from differential hybridization of array probes by polymorphic transcripts (Alberts et al. 2007). In the area of proteomics, Hu et al. (2005) investigated variability in expression of proteins of human

cerebrospinal fluid and Zhang et al. (2006) reported variability in protein expression between 12 human liver samples. These, and other studies on fish (George et al. pers. comm.), identified a very large dynamic range of variation in protein expression.

Experimental design and biological variation

Here we highlight recent findings from studies of variability within both the chemical risk assessment and environmental monitoring scenarios. Hines et al. (pers. comm.) recently conducted an extensive investigation into sources of metabolic variation in marine mussels sampled from the environment. An initial analysis of the entire metabolic dataset showed large variability but the majority of this could be rationalised in terms of season (i.e. month when animal was sampled), sex, species and site effects. This data suggests that information about these parameters is necessary for interpretation of inter-individual biological variation and to potentially reveal pollutant effects.

Similarly, many studies in both fish and rodents have highlighted the strong influence that parameters such as sex and stage of the reproductive cycle have on an individual's transcriptome and proteome, not only in the reproductive tissues but also in other commonly studied tissues such as the brain and liver. In many cases this can be attributed directly to interaction with sex hormone signalling pathways; however, more subtle effects can also be due to cross-talk between nuclear transcription factors. Therefore, it is essential to consider and document both the sex and the stage of the reproductive cycle in

toxicological studies even when assessing chemicals which are not suspected of causing endocrine disruption. While differences in the gonad are to be expected, extensive sexually dimorphic gene expression is also found in somatic tissues in rodents (Boyle and Craft 2000; Waxman and O'Connor 2006; Yang et al. 2006), zebrafish (Robison et al. 2008; Santos et al. 2008; Sreenivasan et al. 2008) and *M. edulis* (Brown et al. 2006) and is likely to occur across species and in particular in organisms developing as gonochoristic males or females. Environmental factors such as temperature and hypoxia have also been shown to strongly influence the transcriptomic and proteomic profiles in poikilotherms. Indeed the expression of several enzymes and proteins which activate and detoxify important chemical toxins, as well as those involved in RNA processing, translation initiation, mitochondrial metabolism, proteasomal function, and essential fatty acid synthesis in fish show a clear temperature dependence (Gracey et al. 2004). Thus the response to chemical exposure may be significantly modulated by the ambient temperature and this must be taken in to account in field sampling. The general recommendation of European legislative studies of chemical impacts is to standardise the sampling season to minimise effects of temperature and the nutritional status and to utilise animals (usually males) when they are either sexually immature or gonadally quiescent.

Key components that require attention to minimise variability, as discussed above, are sampling processes, sample isolation and storage, preparation of samples for analyses, numbers of replicates, design of platforms and methods of normalisation. Particularly important are study design and statistical analyses. Only with biological replicates is it

possible to apply statistical tests. Statistical justification and formula presented by Kendzierski et al. (2003) and Peng et al. (2003) can be used to calculate the number of biological samples and pools needed for an appropriately powered analysis. For omics based studies the challenge in power analysis is how to determine the appropriate variance since tens of thousands of elements (i.e. genes, proteins or metabolites) are measured simultaneously and each one possesses its own variance. One option is to use a pooled variance, such as the value estimated from PCA or other error pooling algorithms. The second option is to calculate the variance for each element and take the n^{th} percentile as the value to use. Usually, detection of small changes for elements with large variances is unlikely. By adjusting the variance value and the effect size (i.e. fold change), power analysis can be tailored to choose the sample size that fits the purpose and significance of an experiment. Power analysis requires a pre-existing data set that resembles the proposed project as much as possible in order to derive an accurately estimated variance. The perfect data set comes from a pilot study for the proposed project. When that is impossible, we recommend matching the biological aspects (e.g. species and tissues) before matching the technical aspects (e.g. technologies and platforms), assuming the technical variance is smaller than the biological variance.

References

Alberts R, Terpstra P, Li Y, Breitling R, Nap JP, Jansen RC. 2007. Sequence polymorphisms cause many false cis eQTLs. PLoS ONE 2(7): e622.

Arikawa E, Sun Y, Wang J, Zhou Q, Ning B, Dial SL, et al. 2008. Cross-platform comparison of SYBR Green real-time PCR with TaqMan PCR, microarrays and other gene expression measurement technologies evaluated in the MicroArray Quality Control (MAQC) study. *BMC Genomics* 9: 328.

Baker SC, Bauer SR, Beyer RP, Brenton JD, Bromley B, Burrill J, et al. 2005. The External RNA Controls Consortium: a progress report. *Nat Methods* 2(10): 731-734.

Bammler T, Beyer RP, Bhattacharya S, Boorman GA, Boyles A, Bradford BU, et al. 2005. Standardizing global gene expression analysis between laboratories and across platforms. *Nat Methods* 2(5): 351-356

Bertram HC, Malmendal A, Petersen BO, Madsen JC, Pedersen H, Nielsen NC, et al. 2007. Effect of magnetic field strength on NMR-based metabonomic human urine data. Comparative study of 250, 400, 500, and 800 MHz. *Anal Chem* 79(18): 7110-7115.

Beyer RP, Fry RC, Lasarev MR, McConnachie LA, Meira LB, Palmer VS, et al. 2007. Multicenter study of acetaminophen hepatotoxicity reveals the importance of biological endpoints in genomic analyses. *Toxicol Sci* 99(1): 326-337.

Boedigheimer MJ, Wolfinger RD, Bass MB, Bushel PR, Chou JW, Cooper M, et al. 2008. Sources of variation in baseline gene expression levels from toxicogenomics study control animals across multiple laboratories. *BMC Genomics* 9: 285.

Boyle SP, Craft JA. 2000. Gender-specific metabolism of benz[a]anthracene in hepatic microsomes from Long-Evans and Hooded Lister rats. *Chem Biol Interact* 125(3): 209-220

Brown M, Davies IM, Moffat CF, Craft JA. 2006. Application of SSH and a macroarray to investigate altered gene expression in *Mytilus edulis* in response to exposure to benzo[a]pyrene. *Mar Environ Res* 62 Suppl: S128-135.

Canales RD, Luo Y, Willey JC, Austermiller B, Barbacioru CC, Boysen C, et al. 2006. Evaluation of DNA microarray results with quantitative gene expression platforms. *Nat Biotechnol* 24(9): 1115-1122.

Chen JJ, Hsueh HM, Delongchamp RR, Lin CJ, Tsai CA. 2007. Reproducibility of microarray data: a further analysis of microarray quality control (MAQC) data. *BMC Bioinformatics* 8: 412.

de Koning DJ, Jaffrezic F, Lund MS, Watson M, Channing C, Hulsege I, et al. 2007. The EADGENE Microarray Data Analysis Workshop. *Genet Sel Evol* 39(6): 621-631.

Fielden MR, Nie A, McMillian M, Elangbam CS, Trela BA, Yang Y, et al. 2008. Interlaboratory evaluation of genomic signatures for predicting carcinogenicity in the rat. *Toxicol Sci* 103(1): 28-34.

Fusco JC, Tong W, Shi L. 2007. QA/QC issues to aid regulatory acceptance of microarray gene expression data. *Environ Mol Mutagen* 48(5): 349-353.

Gebhardt R. 1992. Metabolic zonation of the liver: regulation and implications for liver function. *Pharmacol Ther* 53(3): 275-354.

Gracey AY, Fraser EJ, Li W, Fang Y, Taylor RR, Rogers J, et al. 2004. Coping with cold: An integrative, multitissue analysis of the transcriptome of a poikilothermic vertebrate. *Proc Natl Acad Sci U S A* 101(48): 16970-16975.

Hershey AD, Burdine D, Liu C, Nick TG, Gilbert DL, Glauser TA. 2008. Assessing quality and normalization of microarrays: case studies using neurological genomic data. *Acta Neurol Scand* 118(1): 29-41.

Hines A, Oladiran GS, Bignell JP, Stentiford GD, Viant MR. 2007. Direct sampling of organisms from the field and knowledge of their phenotype: key recommendations for environmental metabolomics. *Environ Sci Technol* 41(9): 3375-3381.

Hu Y, Malone JP, Fagan AM, Townsend RR, Holtzman DM. 2005. Comparative proteomic analysis of intra- and interindividual variation in human cerebrospinal fluid. *Mol Cell Proteomics* 4(12): 2000-2009.

Irizarry RA, Hobbs B, Collin F, Beazer-Barclay YD, Antonellis KJ, Scherf U, et al. 2003. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* 4(2): 249-264.

Kendzierski CM, Zhang Y, Lan H, Attie AD. 2003. The efficiency of pooling mRNA in microarray experiments. *Biostatistics* 4(3): 465-477.

Keun HC, Ebbels TM, Antti H, Bollard ME, Beckonert O, Schlotterbeck G, et al. 2002. Analytical reproducibility in (1)H NMR-based metabonomic urinalysis. *Chem Res Toxicol* 15(11): 1380-1386.

Kohlmann A, Kipps TJ, Rassenti LZ, Downing JR, Shurtleff SA, Mills KI, et al. 2008. An international standardization programme towards the application of gene expression profiling in routine leukaemia diagnostics: the Microarray Innovations in LEukemia study prephase. *Br J Haematol* 142(5): 802-807.

Kuo WP, Liu F, Trimarchi J, Punzo C, Lombardi M, Sarang J, et al. 2006. A sequence-oriented comparison of gene expression measurements across different hybridization-based technologies. *Nat Biotechnol* 24(7): 832-840

Mattes WB. 2008. Public consortium efforts in toxicogenomics. *Methods Mol Biol* 460: 221-238.

Morrison N, Bearden D, Bundy JG, Collette T, Currie F, Davey P et al. 2007. Standard reporting requirements for biological samples in metabolomics experiments: environmental context. *Metabolomics* 3(3): 203-210.

Oinonen T, Lindros KO. 1998. Zonation of hepatic cytochrome P-450 expression and regulation. *Biochem J* 329 (Pt 1): 17-35.

Patterson TA, Lobenhofer EK, Fulmer-Smentek SB, Collins PJ, Chu TM, Bao W, et al. 2006. Performance comparison of one-color and two-color platforms within the MicroArray Quality Control (MAQC) project. *Nat Biotechnol* 24(9): 1140-1150.

Peng X, Wood CL, Blalock EM, Chen KC, Landfield PW, Stromberg AJ. 2003. Statistical implications of pooling RNA samples for microarray experiments. *BMC Bioinformatics* 4: 26.

Robison BD, Drew RE, Murdoch GK, Powell M, Rodnick KJ, Settles M, et al. 2008. Sexual dimorphism in hepatic gene expression and the response to dietary carbohydrate manipulation in the zebrafish (*Danio rerio*). *Comparative Biochemistry and Physiology Part D: Genomics and Proteomics* 3(2): 141-154.

Santos EM, Kille P, Workman VL, Paull GC, Tyler CR. 2008. Sexually dimorphic gene expression in the brains of mature zebrafish. *Comp Biochem Physiol A Mol Integr Physiol* 149(3): 314-324.

Shi L, Reid LH, Jones WD, Shippy R, Warrington JA, Baker SC, et al. 2006. The MicroArray Quality Control (MAQC) project shows inter- and intraplatform reproducibility of gene expression measurements. *Nat Biotechnol* 24(9): 1151-1161.

Sreenivasan R, Cai M, Bartfai R, Wang X, Christoffels A, Orban L. 2008. Transcriptomic analyses reveal novel genes with sexually dimorphic expression in the zebrafish gonad and brain. *PLoS ONE* 3(3): e1791.

Thompson KL, Pine PS, Rosenzweig BA, Turpaz Y, Retief J. 2007. Characterization of the effect of sample quality on high density oligonucleotide microarray data using progressively degraded rat liver RNA. *BMC Biotechnol* 7: 57

Thompson KL, Hackett J. 2008. Quality control of microarray assays for toxicogenomic and in vitro diagnostic applications. *Methods Mol Biol* 460: 45-68

Thorpe KL, Hutchinson TH, Hetheridge MJ, Scholze M, Sumpter JP, Tyler CR. 2001. Assessing the biological potency of binary mixtures of environmental estrogens using vitellogenin induction in juvenile rainbow trout (*Oncorhynchus mykiss*). *Environ Sci Technol* 35(12): 2476-2481.

Thorpe KL, Cummings RI, Hutchinson TH, Scholze M, Brighty G, Sumpter JP, et al. 2003. Relative potencies and combination effects of steroidal estrogens in fish. *Environ Sci Technol* 37(6): 1142-1149.

Viant MR, Bearden DW, Bundy JG, Burton IW, Collette TW, Ekman DR, et al. 2009. International NMR-based environmental metabolomics intercomparison exercise. *Environ Sci Technol* 43(1): 219-225.

Walker MS, Hughes TA. 2008. Messenger RNA expression profiling using DNA microarray technology: diagnostic tool, scientific analysis or un-interpretable data? *Int J Mol Med* 21(1): 13-17.

Waxman DJ, O'Connor C. 2006. Growth hormone regulation of sex-dependent liver gene expression. *Mol Endocrinol* 20(11): 2613-2629.

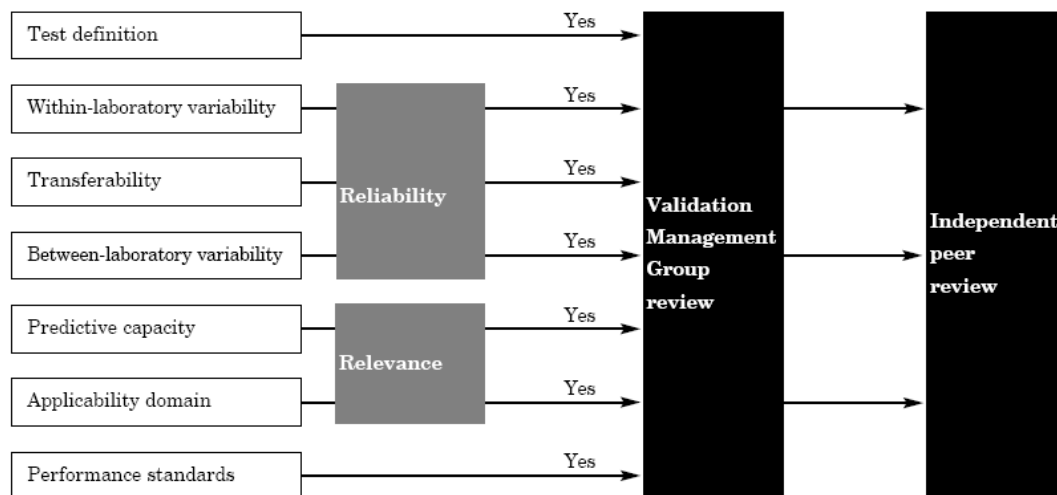
Williams RB, Chan EK, Cowley MJ, Little PF. 2007. The influence of genetic variation on gene expression. *Genome Res* 17(12): 1707-1716.

Yang X, Schadt EE, Wang S, Wang H, Arnold AP, Ingram-Drake L, et al. 2006. Tissue-specific expression and regulation of sexually dimorphic genes in mice. *Genome Res* 16(8): 995-1004.

Zhang X, Guo Y, Song Y, Sun W, Yu C, Zhao X, et al. 2006. Proteomic analysis of individual variation in normal livers of human beings using difference gel electrophoresis. *Proteomics* 6(19): 5260-5268.

Supplemental Material, Figure 1.

The modular approach for applying the ECVAM principles on test validity



A "yes" indicates that the appropriate information for the module is adequate for entrance into the peer-review process. All seven modules have to be satisfactorily completed, as judged by the Validation Management Group, before a method can enter the peer-review process.

ECVAM = European Centre for the Validation of Alternative Methods.

Reproduced from Hartung et al (2004), with permission.

Supplemental Material, Table 1: Key examples of the successful use of toxicogenomics data in ecotoxicology.

A. Prospective Studies

APPLICATION	NATURE OF EXPOSURE	APPROACHES	SIGNIFICANCE OF RESULTS	REFERENCES
Chemical Signatures	Diet, static renewal, semi-static renewal or <i>ip</i> injection	Microarray	Demonstrates chemical class specific gene expression, and/or chemical specific signatures	Benninghoff and Williams 2008; Larkin et al. 2003; Hamadeh et al. 2002; McMillian et al 2005; Moens et al. 2006; Hook et al. 2006; Ellinger-Ziegelbauer et al. 2008
Mixtures	Hepatocytes in vitro	Microarray	Demonstrates that mixtures show attenuated biomarker responses compared to those of the chemical signatures of classical environmental toxicants	Finne et al. 2007
Endocrine Disruption Pathways	Flow-through	Microarray	Links reduced fecundity/population decline to perturbations in endocrine pathways, vitellogenin, steroid concentrations, gonad weight, and specific perturbations in gene expression	Miller et al. 2007; Ankley et al. 2008; Villeneuve et al. 2007; Hoffmann et al. 2008
Pathways of Toxicity	Injection	Microarray	Demonstrates that TCDD perturbs fin regeneration by impacting the expression of genes involved in extracellular matrix composition and cellular differentiation	Andreasen et al. 2006
Pathways of Toxicity	Static exposure	Microarray	Demonstrates heart-specific mechanisms of AhR/TCDD-mediated toxicity	Handley-Goldstone et al.

				2005; Carney et al. 2006
Pathways of Toxicity/ Biomarkers of Effect and Temporal Changes	Flow-through	Microarray	Used cDNA arrays and Q-PCR to identify potential indicators of thyroid axis and metamorphosis disruption in frogs	Helbing et al. 2007a; Helbing et al. 2007b
Dose-Response, Adaptive/Toxic Response	Static renewal	Microarray	Links gene expression changes to reduced growth.	Roling et al. 2006
Temporal and Adaptive Changes	Injection	Microarray	Demonstrates time-dependent adaptive changes prior to toxicity following TCDD treatment	Volz et al. 2006
Screening of Emerging Chemicals; Predicting Adverse Outcomes	Flow-through	Microarray	Uses microarrays to determine the potential mechanisms of PFOA toxicity	Wei et al. 2008
Emerging Chemicals	Static	Microarray	Uses microarrays to demonstrate differential effects of nanoparticles and their constituents	Griffitt et al. 2007
Inter-laboratory Comparisons	Flow-through, other	Microarray	Meta-analysis of data from environmental estrogen exposure that produced new, sensitive biomarkers of exposure	Gunnarsson et al. 2007
Adaptive/Toxic Response	Static renewal	Microarray	Demonstrates metal specific gene expression in response to copper, cadmium, and zinc. Proposes novel modes of toxicant action.	Poynton et al. 2007
Computational	Flow-through	Microarray	Assesses sources of variation in fish microarray experiments. Chemical class prediction using bioinformatic classification software such as Support vector machines	Wang et al. 2008a, Wang et al. 2008b
Chemical Signatures/ Adaptive Response	Flow-through	Metabolomics	Demonstrates compensatory mechanisms and adaptive recovery from an estrogen. Also demonstrates the potential of metabolomics in	Ekman et al. 2007

			ecotoxicology.	
Temporal Changes	Flow-through	Proteomics	Provides early response indicators to thyroid hormones in frogs	Domanski and Helbing 2007
Chemical signatures	laboratory	Metabolomics	A mechanism of action from combination of NMR metabolite profiling and neural network classification	Ott et al 2003

B. Diagnostic Studies

APPLICATION	NATURE OF EXPOSURE	APPROACHES	SIGNIFICANCE OF RESULTS	REFERENCES
Population Genetics	None/unknown field	Microarray	Demonstrate variation in gene expression and potential difficulties of using particular genes as biomarkers	Oleksiak et al. 2002; Oleksiak et al. 2005
Sample Monitoring	Wastewater treatment	Microarray	Effect of Mixtures; endocrine disruptors and no-endocrine disruptors	Filby et al. 2007b; Garcia-Reyero et al. 2008; Filby et al. 2007c
Sample Monitoring	Field	qRT-PCR	Demonstrates that seal thyroid hormones are sensitive to disruption by pollutant stress	Tabuchi et al. 2006
Population Genetics	Field	Microarray	Demonstrates differential adaptation to distinctly different field sites	Larsen et al. 2007
Chemical Remediation	Field	Microarray	Demonstrates the utility of microarrays as an additional weight-of-evidence approach to monitor remediation at a polluted site	Roling et al. 2007
Site Monitoring and computational	Field combined with laboratory	Microarray, Differential display	Demonstrates differential expression at polluted and reference sites, and demonstrates distinct differences between males and females in the pollutant responses. Prediction of environmental source of fish by stress gene responses	Williams et al. 2003; Falciani et al. 2008, Meyer et al. 2005
Site Monitoring	Field	Subtractive hybridization, Differential display	Demonstrates differential expression at polluted and references site. Links to lab work to field sampling	Maples and Bain 2004; Roling et al. 2004
Computational	Field	Metabolomics	Improved discrimination between sample	Parsons et al. 2007

			classes compared to unscaled, autoscaled or Pareto scaled data. Leads to better site classification	
Site monitoring	Field	Metabolomics	Demonstrates differential metabolic fingerprints in earthworms (<i>Lumbricus rubellus</i>) from sites with differing metal contamination.	Bundy et al. 2007
Site monitoring	Field	Metabolomics	Demonstrates differential metabolic fingerprints in marine mussels (<i>Mytilus edulis</i>) at polluted and reference sites.	Viant, unpublished

REFERENCES

Andreasen EA, Mathew LK, Tanguay RL. 2006. Regenerative growth is impacted by TCDD: gene expression analysis reveals extracellular matrix modulation. *Toxicol Sci* 92(1): 254-269.

Ankley GT, Miller DH, Jensen KM, Villeneuve DL, Martinovic D. 2008. Relationship of plasma sex steroid concentrations in female fathead minnows to reproductive success and population status. *Aquat Toxicol* 88(1): 69-74.

Benninghoff AD, Williams DE. 2008. Identification of a transcriptional fingerprint of estrogen exposure in rainbow trout liver. *Toxicol Sci* 101(1): 65-80.

Bundy JG, Keun HC, Sidhu JK, Spurgeon DJ, Svendsen C, Kille P, et al. 2007. Metabolic profile biomarkers of metal contamination in a sentinel terrestrial species are applicable across multiple sites. *Environ Sci Technol* 41(12): 4458-4464.

Carney SA, Chen J, Burns CG, Xiong KM, Peterson RE, Heideman W. 2006. Aryl hydrocarbon receptor activation produces heart-specific transcriptional and toxic responses in developing zebrafish. *Mol Pharmacol* 70(2): 549-561.

Domanski D, Helbing CC. 2007. Analysis of the *Rana catesbeiana* tadpole tail fin proteome and phosphoproteome during T3-induced apoptosis: identification of a novel type I keratin. *BMC Dev Biol* 7: 94.

Ekman DR, Teng Q, Jensen KM, Martinovic D, Villeneuve DL, Ankley GT, et al. 2007. NMR analysis of male fathead minnow urinary metabolites: a potential approach for studying impacts of chemical exposures. *Aquat Toxicol* 85(2): 104-112.

Ellinger-Ziegelbauer H, Gmuender H, Bandenburg A, Ahr HJ. 2008. Prediction of a carcinogenic potential of rat hepatocarcinogens using toxicogenomics analysis of short-term in vivo studies. *Mutat Res* 637(1-2): 23-39.

Falciani F, Diab AM, Sabine V, Williams TD, Ortega F, George SG, et al. 2008. Hepatic transcriptomic profiles of European flounder (*Platichthys flesus*) from field sites and computational approaches to predict site from stress gene responses following exposure to model toxicants. *Aquat Toxicol* 90(2): 92-101

Filby AL, Santos EM, Thorpe KL, Maack G, Tyler CR. 2007b. Gene expression profiling for understanding chemical causation of biological effects for complex mixtures: a case study on estrogens. *Environ Sci Technol* 41(23): 8187-8194.

Filby AL, Thorpe KL, Maack G, Tyler CR. 2007. Gene expression profiles revealing the mechanisms of anti-androgen- and estrogen-induced feminization in fish. *Aquat Toxicol* 81(2): 219-231.

Finne EF, Cooper GA, Koop BF, Hylland K, Tollefsen KE. 2007. Toxicogenomic responses in rainbow trout (*Oncorhynchus mykiss*) hepatocytes exposed to model chemicals and a synthetic mixture. *Aquat Toxicol* 81(3): 293-303.

Garcia-Reyero N, Adelman I, Liu L, Denslow N. 2008. Gene expression profiles of fathead minnows exposed to surface waters above and below a sewage treatment plant in Minnesota. *Mar Environ Res* 66(1): 134-136.

Griffitt RJ, Weil R, Hyndman KA, Denslow ND, Powers K, Taylor D, et al. 2007. Exposure to copper nanoparticles causes gill injury and acute lethality in zebrafish (*Danio rerio*). *Environ Sci Technol* 41(23): 8178-8186.

Gunnarsson L, Kristiansson E, Forlin L, Nerman O, Larsson DG. 2007. Sensitive and robust gene expression changes in fish exposed to estrogen--a microarray approach. *BMC Genomics* 8: 149.

Hamadeh HK, Bushel PR, Jayadev S, Martin K, DiSorbo O, Sieber S, et al. 2002. Gene expression analysis reveals chemical-specific profiles. *Toxicol Sci* 67(2): 219-231.

Handley-Goldstone HM, Grow MW, Stegeman JJ. 2005. Cardiovascular gene expression profiles of dioxin exposure in zebrafish embryos. *Toxicol Sci* 85(1): 683-693.

Helbing CC, Bailey CM, Ji L, Gunderson MP, Zhang F, Veldhoen N, et al. 2007. Identification of gene expression indicators for thyroid axis disruption in a *Xenopus laevis* metamorphosis screening assay. Part 1. Effects on the brain. *Aquat Toxicol* 82(4): 227-241.

Helbing CC, Ji L, Bailey CM, Veldhoen N, Zhang F, Holcombe GW, et al. 2007. Identification of gene expression indicators for thyroid axis disruption in a *Xenopus laevis* metamorphosis screening assay. Part 2. Effects on the tail and hindlimb. *Aquat Toxicol* 82(4): 215-226.

Hoffmann JL, Thomason RG, Lee DM, Brill JL, Price BB, Carr GJ, et al. 2008. Hepatic gene expression profiling using genechips in zebrafish exposed to 17alpha-methyldihydrotestosterone. *Aquat Toxicol* 87(2): 69-80.

Hook SE, Skillman AD, Small JA, Schultz IR. 2006. Gene expression patterns in rainbow trout, *Oncorhynchus mykiss*, exposed to a suite of model toxicants. *Aquat Toxicol* 77(4): 372-385.

Larkin P, Folmar LC, Hemmer MJ, Poston AJ, Denslow ND. 2003. Expression profiling of estrogenic compounds using a sheepshead minnow cDNA macroarray. *EHP Toxicogenomics* 111(1T): 29-36.

Larsen PF, Nielsen EE, Williams TD, Hemmer-Hansen J, Chipman JK, Kruhoffer M, et al. 2007. Adaptive differences in gene expression in European flounder (*Platichthys flesus*). *Mol Ecol* 16(22): 4674-4683.

Maples NL, Bain LJ. 2004. Trivalent chromium alters gene expression in the mummichog (*Fundulus heteroclitus*). *Environ Toxicol Chem* 23(3): 626-631.

McMillian M, Nie A, Parker JB, Leone A, Kemmerer M, Bryant S, et al. 2005. Drug-induced oxidative stress in rat liver from a toxicogenomics perspective. *Toxicol Appl Pharmacol* 207(2 Suppl): 171-178.

Meyer JN, Volz DC, Freedman JH, Giulio RT. 2005. Differential display of hepatic mRNA from killifish (*Fundulus heteroclitus*) inhabiting a Superfund estuary. *Aquat Toxicol* 73(4): 327-341.

Miller DH, Jensen KM, Villeneuve DL, Kahl MD, Makynen EA, Durhan EJ, et al. 2007. Linkage of biochemical responses to population-level effects: a case study with vitellogenin in the fathead minnow (*Pimephales promelas*). *Environ Toxicol Chem* 26(3): 521-527.

Moens LN, van der Ven K, Van Remortel P, Del-Favero J, De Coen WM. 2006. Expression profiling of endocrine-disrupting compounds using a customized *Cyprinus carpio* cDNA microarray. *Toxicol Sci* 93(2): 298-310.

Oleksiak MF, Churchill GA, Crawford DL. 2002. Variation in gene expression within and among natural populations. *Nat Genet* 32(2): 261-266.

Oleksiak MF, Roach JL, Crawford DL. 2005. Natural variation in cardiac metabolism and gene expression in *Fundulus heteroclitus*. *Nat Genet* 37(1): 67-72.

Ott KH, Aranibar N, Singh B, Stockton GW. Metabonomics classifies pathways affected by bioactive compounds. Artificial neural network classification of NMR spectra of plant extracts. *Phytochemistry* 2003; 62: 971-985.

Parsons HM, Ludwig C, Gunther UL, Viant MR. 2007. Improved classification accuracy in 1- and 2-dimensional NMR metabolomics data using the variance stabilising generalised logarithm transformation. *BMC Bioinformatics* 8: 234

Poynton HC, Varshavsky JR, Chang B, Cavigiolio G, Chan S, Holman PS, et al. 2007. *Daphnia magna* ecotoxicogenomics provides mechanistic insights into metal toxicity. *Environ Sci Technol* 41(3): 1044-1050.

Roling JA, Bain LJ, Baldwin WS. 2004. Differential gene expression in mummichogs (*Fundulus heteroclitus*) following treatment with pyrene: comparison to a creosote contaminated site. *Mar Environ Res* 57(5): 377-395.

Roling JA, Bain LJ, Gardea-Torresdey J, Bader J, Baldwin WS. 2006. Hexavalent chromium reduces larval growth and alters gene expression in mummichog (*Fundulus heteroclitus*). *Environ Toxicol Chem* 25(10): 2725-2733.

Roling JA, Bain LJ, Gardea-Torresdey J, Key PB, Baldwin WS. 2007. Using mummichog (*Fundulus heteroclitus*) arrays to monitor the effectiveness of remediation at a superfund site in Charleston, South Carolina, U.S.A. *Environ Toxicol Chem* 26(6): 1205-1213.

Tabuchi M, Veldhoen N, Dangerfield N, Jeffries S, Helbing CC, Ross PS. 2006. PCB-related alteration of thyroid hormones and thyroid hormone receptor gene expression in free-ranging harbor seals (*Phoca vitulina*). *Environ Health Perspect* 114(7): 1024-1031.

Villeneuve DL, Larkin P, Knoebl I, Miracle AL, Kahl MD, Jensen KM, et al. 2007. A graphical systems model to facilitate hypothesis-driven ecotoxicogenomics research on the teleost brain-pituitary-gonadal axis. *Environ Sci Technol* 41(1): 321-330.

Volz DC, Hinton DE, Law JM, Kullman SW. 2006. Dynamic gene expression changes precede dioxin-induced liver pathogenesis in medaka fish. *Toxicol Sci* 89(2): 524-534.

Wang RL, Bencic D, Biales A, Lattier D, Kostich M, Villeneuve D, et al. 2008. DNA microarray-based ecotoxicological biomarker discovery in a small fish model species. *Environ Toxicol Chem* 27(3): 664-675.

Wang RL, Biales A, Bencic D, Lattier D, Kostich M, Villeneuve D, Ankley GT, Lazorchak J, Toth G. 2008. DNA microarray application in ecotoxicology: experimental design, microarray scanning, and factors affecting transcriptional profiles in a small fish species. *Environ Toxicol Chem*; 27: 652-663.

Wei Y, Liu Y, Wang J, Tao Y, Dai J. 2008. Toxicogenomic analysis of the hepatic effects of perfluorooctanoic acid on rare minnows (*Gobiocypris rarus*). *Toxicol Appl Pharmacol* 226(3): 285-297.

Williams TD, Gensberg K, Minchin SD, Chipman JK. 2003. A DNA expression array to detect toxic stress response in European flounder (*Platichthys flesus*). *Aquat Toxicol* 65(2): 141-157.